High-speed Developments in Avian Genomics

High-Speed Developments in Avian Genomics

CAMILLE BONNEAUD, JOAN BURNSIDE, AND SCOTT V. EDWARDS

Until recently, definitions of avian genome structure and function were based solely on our knowledge of the chicken genome. The expansion of genomic studies to include nonmodel avian species allows us not only to refine those definitions but also to begin collecting the necessary resources to initiate a truly ecological genomics of birds. In this article we review new genomic technologies that will speed up the investigation of avian genome function. The streamlined nature of avian genomes implies that large-scale transcriptional analyses, studies of the role of regulatory elements and of developmental genes, and even the annotation of avian genomes will yield interesting surprises. We review promising methods used to investigate genome evolution in birds as well as the means by which to integrate functional genomics approaches and transcriptional profiling information into ecological and evolutionary studies.

Keywords: avian genome, genome sequencing, candidate genes, transcription profiles

Just as the polymerase chain reaction leveled the genetic playing field at the end of the 20th century by providing easy access to the genes of all organisms, so the 21st century promises to sweep away the technological privileges of classical model organisms and democratize genomic exploration. Excursions into the genomes of nonmodel species are now made possible by our growing ability to sequence whole genomes. In addition, evolutionary biologists are increasingly making use of molecular tools that do not require prior genetic information of the study species, or that are vastly facilitated by information from closely related species. By building on the substantial genetic database developed for model organisms, we can begin to explore the genomic landscapes of outbred species, often sampled from natural populations, and take a detailed look at the evolutionary process.

This revolution is particularly exciting to those whose study species are genomically underrepresented. For example, until recently, the chicken (*Gallus gallus*) was the only bird in this select community of study species (ICGSC 2004). Yet the chicken genome has revealed striking differences from the genomic structure and functioning found in mammals (for a review, see Ellegren [2007]). In addition, since birds and mammals are thought to have diverged approximately 310 million years ago (Hedges 2002), genomic comparisons between both groups should yield important information regarding genome evolution (Hardison 2003, Ellegren 2005). The genome of the zebra finch (*Taeniopygia guttata*), a species widely used in neurogenetic and developmental studies of avian song, is currently being sequenced and will be the second bird to join the ranks of so-called model organisms. Extending genomic information from these species to the remaining 10,000 species of birds will present exciting challenges and opportunities.

Here we review some of the genomic techniques currently used to explore the avian genome, with a particular focus on studies with an evolutionary approach. Our goal is not to cover all studies seeking to understand the function and evolution of avian genes, nor to exhaustively review the emerging picture of avian genome evolution (Edwards et al. 2005, Ellegren 2005). Rather, we wish to highlight structural genomic features that might facilitate functional genomic studies in birds, and to give evolutionary biologists an overview of emerging means to study genomic adaptation.

The avian genome—a streamlined genome

A quick search on the Animal Genome Size Database (Gregory 2007) is enough to show that birds are positioned at the low end of the spectrum of genome size among vertebrates. Although the smaller genome sizes of birds are obvious from casual inspection, Waltari and Edwards (2002) were the first to show that this distinction withstands the scrutiny of phylogenetically controlled comparisons. This and recent additional comparative analyses (Organ et al. 2007) have shown that long-term genome size evolution in birds is characterized by strong stabilizing selection, keeping modern avian genomes within a narrow size range, as well as by stasis, given recent findings that nonavian dinosaurs originating more than 200 million years ago very likely possessed genomes as small as those of modern birds.

Camille Bonneaud is a postdoctoral research fellow, and Scott V. Edwards (e-mail: sedwards@fas.harvard.edu) is a professor in evolutionary biology, in the Department of Organismic and Evolutionary Biology at Harvard University in Cambridge, Massachusetts. Joan Burnside is a professor in the College of Agriculture and Natural Resources at the University of Delaware in Newark.

The current record for the smallest avian genome is held by the common pheasant, *Phasianus colchicus,* with a size of 0.97 picogram (pg; genome size is given as the mass of DNA [deoxyribonucleic acid] per haploid nucleus), whereas the ostrich, *Struthio camelus,* wins the prize for the largest avian genome with a size of 2.16 pg (Gregory 2007), suggesting a possible role for flight in genome size reduction in birds (Hughes 2000). Indeed, small genomes were originally thought to facilitate the metabolic demands of flight by enabling the evolution of small cell sizes. However, the small genomes of birds are now known to have originated deep within the dinosaurian roots of modern birds long before the origin of flight, perhaps as a means of accommodating other metabolic needs (Organ et al. 2007). Regardless, the decreased size of avian genomes not only implies a restructured genomic platform for gene regulation and the organization of noncoding elements but also suggests that studying genome processes in birds might be greatly facilitated by this streamlining.

Our understanding of the basics of gene regulation, transcription, and indeed the definition of the gene itself are undergoing radical revision (Gerstein et al. 2007, Henikoff 2007). It therefore stands to reason that birds might facilitate functional genomics studies because their genomes, having jettisoned gigabases of retroelements, complex repetitive regions, and noncoding DNA, presumably increase the signal-to-noise ratio of genome functioning over that in mammals. On the other hand, the higher morphological complexity of mammals, in comparison with birds (Wyles et al. 1983), may have been facilitated by their more complex genomes.

The small genomes of birds have additional consequences for genome function, dynamics, and accessibility. For example, the chicken MHC (major histocompatibility complex) B locus, a critical gateway to the adaptive immune response, has long been suggested to have been structurally and functionally streamlined (Kaufman et al. 1999). However, comparative studies of the MHC in a number of bird lineages (Shaw et al. 2007, Strand et al. 2007) and the increasing complexity of the structure and functions of the *Rfp-Y* region (Shiina et al. 2007) suggest that the minimal essential MHC hypothesis may need to be relaxed or may be taxonomically restricted (LePage et al. 2000, Iglesias et al. 2003). A global scan for gene expression in the B-complex and *Rfp-Y* regions would greatly increase our understanding of the functional complexity of this immunologically important region.

Additional correlates of small genomes in birds include a high degree of long-term karyotypic conservation (reviewed in Burt [2002]); a drastic reduction in the incidence and complexity of retroelements and microsatellites (Primmer et al. 1997, Hughes and Piontkivska 2005); a high rate of recombination, particularly in microchromosomes, although precise estimates vary depending on method and scale of measurement (Burt 2002, Dawson et al. 2007, Backström et al. 2008); and increased heterogeneity in rates and patterns of nucleotide substitution across the genome (Axelsson et al. 2005, Webster et al. 2006). The complex interplay between genome size, microchromosomes, and genome dynamics is challenging to disentangle, and some genomic features, such as reduced numbers of microsatellites and conservative modes of evolution, might be holdovers from genomic trends established in reptile and amniote ancestors (Shedlock et al. 2007). Nonetheless, the intriguing and simplifying structural features of avian genomes are crucial for establishing the minimal genomic toolkit required for efficient and fine-grained gene regulation, immunity, and adaptation to changing environments.

Comparing avian genomes, subgenomes, and candidate genes

A first-draft assembly of the genome of a female red jungle fowl (*Gallus gallus*) was published by the International Chicken Genome Sequencing Consortium (ICGSC 2004), and the chicken genome now serves as a genomic and sequence reference for all other birds. Bacterial artificial chromosome (BAC) libraries (box 1) play an increasingly important role in avian comparative genomics (e.g., Edwards et al. 2005, Romanov et al. 2006). Several BAC libraries from birds other than the chicken have been produced in recent years, including those for the zebra finch (*Taeniopygia guttata*) and the California condor (*Gymnogyps californianus*) (Luo et al. 2006, Romanov et al. 2006). Other libraries of interest to avian geneticists are from nonavian reptiles, produced as part of an effort by the National Science Foundation to provide genomic resources for genomically understudied groups (*www.nsf.gov/bio/pubs/awards/bachome.htm, www.sym-bio.com,* Wang et al. 2006a).

Avian genomes and subgenomes. All of these libraries provide a rich platform on which to build a new generation of comparative genomics studies in birds. For example, Romanov and Dodgson (2006) used BAC libraries from the turkey (*Meleagris gallopavo*) and the zebra finch to align several thousand clones of each species to the chicken physical map using so-called overgo probes (box 1). Confirming expectations, their study showed that overgo probes derived from exons were generally more successful in identifying orthologous BACs in other birds than were probes designed from 3' UTRs (untranslated regions) or other regions, and thus helped establish the taxonomic range over which this approach will be effective. Romanov and colleagues (2006) screened their condor BAC library for genes of known positions in the chicken genome, with the goal of pinpointing candidate loci associated with a disease and thereby helping in population management programs. The authors were able to compare the physical map obtained for the condor with that of the chicken, and they found evidence for a high degree of synteny between the two genomes (Romanov et al. 2006).

BAC and plasmid clone libraries (box 1), and of course genome sequences, are also excellent tools for generating large numbers of markers for phylogenetic and biogeographic studies. Traditionally, phylogenetic studies in birds have used mitochondrial genes, such as *cytochrome b* genes, but the number of studies employing multiple nuclear genes in avian phylogeography is growing (Jennings and Edwards 2005,

Box 1. Definitions.

Antisense riboprobe. An RNA (ribonucleic acid) fragment used to probe for a complementary nucleotide sequence of a target mRNA (messenger RNA) or DNA (deoxyribonucleic acid) of a cell.

Bacterial artificial chromosome (BAC) library. A genomic library that comprises hundreds of thousands of clones, each of which contains large (100–200 kilobase) DNA inserts (Miyake and Amemiya 2004). The archiving and storage of these libraries as individual clones in glycerol stocks in 384-well microtiter plates makes them easily available for characterization, manipulation, and DNA sequencing, all of which will play an increasingly important role in avian comparative genomics.

cDNA library. A library composed of cDNA (complementary DNA) derived from all the mRNA expressed in a specific tissue or cell at a specific time.

cDNA microarrays. A large collection of cDNAs immobilized on a glass slide and serving as probes in the hybridization of two fluorescently labeled samples. This technique can be used in gene expression profiling to estimate the relative quantities of mRNA expressed in each of the two samples.

Expressed sequence tag (EST). A partial sequence of a cDNA molecule, usually unique and identifying a specific gene.

In situ *hybridization.* Hybridization of RNA or cDNA that allows the visualization and localization of specific transcripts in tissues or cells positioned on microscopic slides. It is also possible to obtain quantitative information on levels of expression of specific transcripts in specific locations.

Macroarrays. High-density nylon arrays of cDNAs, often enriched for differentially expressed genes, to which RNA from different samples are hybridized. Macroarrays are cheap and reliable but do not offer the quantitative precision of microarrays.

Overgo probes. Pairs of partially overlapping short probes that correspond to conserved regions of genomes can be radioactively labeled to a high specific activity and are useful in identifying orthologous regions from even distantly related genomes (Romanov and Dodgson 2006).

Plasmid clone libraries. Small-insert genomic libraries. Such libraries can be made directly from a genome, to be used for identifying loci containing simple sequence repeats or anonymous loci for phylogeography, or they can be the product of fractionation of larger genomic clones, such as from a BAC library, for full-coverage sequencing.

Retroviral vector. A retrovirus in which part of the viral genome has been replaced with a foreign DNA sequence of interest and which can then be used to introduce the DNA sequence into a genome of infected host cells.

Subtractive suppression hybridization (SSH). The SSH methodology permits the construction of a cDNA library enriched for cDNAs differentially expressed between experimental samples and controls, usually in a macroarray format. Although the SSH approach is technically straightforward and inexpensive, and it can identify genes that are up- or down-regulated and easily applied to nonmodel species, it is not quantitative, nor is it as sensitive as microarrays are for assaying gene expression.

Bensch et al. 2006). Large-scale sequence scanning of BAC-ends, or of complete or partial BAC clones such as those available in Genbank for a variety of nonmodel species, provides abundant markers for studies of the species from which the library was made, as well as closely related species. Many such markers will be from noncoding regions, but in our experience, such markers still have surprisingly wide taxonomic utility. A recent study (Backström et al. 2008) fully exploited the chicken genome for developing markers for phylogenetic and phylogeographic studies by designing 242 primer pairs in conserved exons flanking introns of suitable size for easy amplification of diverse species. Aside from numerous important insights into the comparative genomics of the five target species investigated for primer amplification, this study will undoubtedly usher in a host of large-scale studies relying on these primers for phylogenetic questions.

Several innovative technologies for sequencing genomic DNA or full-length cDNA (complementary DNA) have recently emerged, which avoid the expense, complication, and biases associated with traditional clone-based sequencing by using direct amplification of templates from DNA. Examples of "deep sequencing" technologies include 454 Life Sciences (figure 1), Illumina/Solexa's Sequencing by Synthesis, and Applied Biosystems SOLiD system. The size of the sequence reads makes these approaches ideal for whole genome resequencing, allowing a novel genome sequence to be compared and assembled with help of a reference sequence (Bentley 2006). But given the online availability of passerine and chicken genomes, these approaches will also be useful for characterization of large genomic regions and transcriptomes of species that do not have sequenced genomes. The application of these new technologies to other birds offers great possibilities for rapid advances in bird comparative genomics.

Candidate gene approaches. Many studies have used candidate gene approaches inspired from model organisms to identify genes and understand processes occurring in nonmodel species. BAC libraries again provide an easy inroad for studying candidate genes of interest to avian biology. For example, Luo and colleagues (2006) recently screened a zebra finch BAC

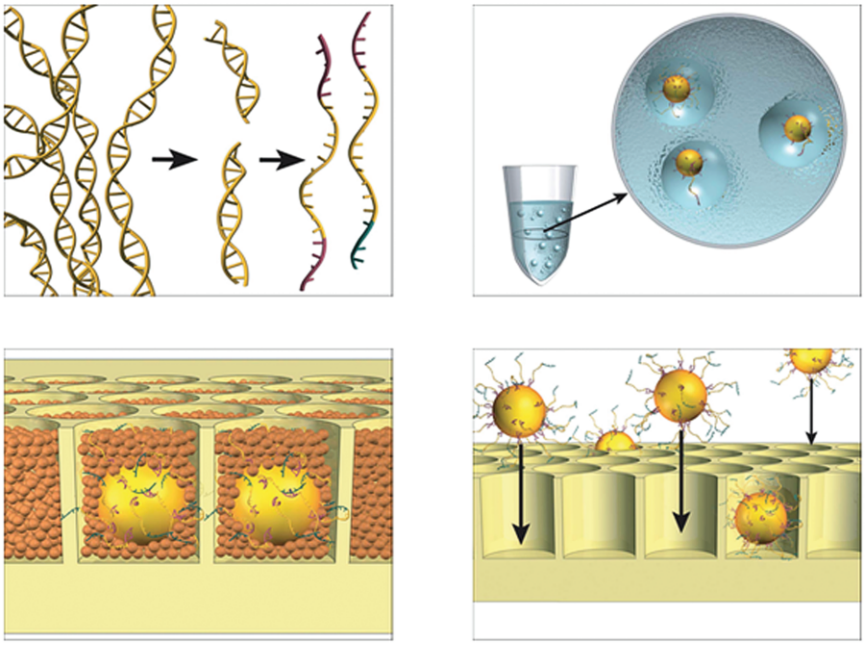

Figure 1. The technology developed by 454 Life Science allows the parallel sequencing of millions of base pairs from millions of DNA fragments, within only a few hours (www.454.com). Large DNA sequences, such as genomic DNA and bacterial artificial chromosomes, are first fractioned into 300 to 800 base pair fragments (top left). Shorter sequences, for example, noncoding RNA and amplicons, do not require fractioning. Fragments are then blunted and short adaptors termed A and B are attached to the 3′ and 5′ ends. Single-stranded fragments containing both A and B are kept in further steps. The B adaptor contains biotin, which fixes the single-stranded template DNA to a bead that is subsequently captured in a droplet of polymerase chain reaction mixture (containing DNA polymerase) in an oil emulsion (top right). Amplification occurs with each droplet so that each bead will end up carrying millions of identical copies of the initial DNA fragment. The beads are then placed in an individual well onto a fiber optic plate called PicoFilterPlate (bottom right). Smaller enzyme beads, containing sulferase and luciferase, are added to each well (bottom left). The PicoFilterPlate is placed in the Genome Sequencer FLX instrument, where the sequencing reaction will subsequently be carried out. The sequencing reagents, including buffer and nucleotides, are flowed across the wells. Nucleotides are flowed sequentially, and each DNA fragment is sequenced in parallel. Each time a nucleotide is added to the template strand, a chemiluminescent signal is produced by the sequencing reaction and recorded by the CCD (charge-coupled device) camera. The signal is then processed and the base sequences analyzed. Yields of more than 200,000 sequences per run, with read lengths in excess of 100 base pairs, generate approximately 20,000,000 base pairs of sequence data. Adapted with permission from Margulies and colleagues (2005).

library and sequenced an open reading frame for the full-length androgen receptor, a gene of vital interest for understanding the developmental genetics of sexual dimorphism and behavior. They also used the BAC library to compare the androgen receptor sequence and flanking sequences from zebra finch with the genome draft sequences from chicken, and demonstrated that this region of the zebra finch genome was highly syntenic to the homologous region of the chicken genome.

The most extensive body of work examining functional genes and changes in the transcription profiles of various candidate genes in birds other than chicken comes from neurogenetic studies of song in passerines, particularly in zebra finches, canaries, and European starlings (for a review, see Clayton [2004]). Much work has also been done on the avian brain (e.g., chicken, Japanese quail, Java sparrow, pigeon, house sparrow, garden warbler, and barn owl), using candidate gene expression profiles to identify genes involved in the circadian clock (e.g., Helfer et al. 2006).

The candidate gene approach in evolutionary investigations of birds includes studies of MHC genes in passerines (e.g., Richardson et al. 2005) and studies examining variation in plumage melanism under the control of the melanocortin-1 receptor (*MC1R*) locus (e.g., Mundy 2005, Nadeau et al. 2007). A focus on the functional and evolutionary consequences of both protein and noncoding variation is beginning to emerge in investigations of the genetic basis of avian behavioral and morphological traits.

Behavioral syndromes (Sih et al. 2004) or animal personalities (Dall et al. 2004) are thought to influence life history traits and fitness, and variation in behavior may be maintained by natural selection (for a review, see van Oers et al. [2005]). For instance, risks of feather-pecking in a flock, which reduce the victim's fitness without any obvious direct benefit to the assailant, was shown to increase severely in chickens expressing a wild recessive allele at a gene controlling feather melanization (*PMEL17*) (Keeling et al. 2004).

Great tits (*Parus major*) are becoming models in the exploration of both proximal and ultimate factors shaping animal personalities (Fidler et al. 2007). In a recent investigation, Fidler and coworkers (2007) sought associations between gene polymorphisms and personality variation in great tits. They focused on a cDNA sequence orthologous to the human dopamine receptor D4 (*Drd4*), reported to be linked to novelty-seeking behavior in humans (reviewed in Kluger et al. 2002). A single nucleotide polymorphism in *Drd4* was found to diverge significantly between two lines of great tit selected for different early exploratory behavior and was also discovered to exhibit significant association with

different levels of early exploratory behavior in free-living and unselected individuals (Fidler et al. 2007). The link between behavioral traits and variation in the *Drd4* gene that was detected in both humans and great tits supports the idea that a given gene may a have similar function in taxonomically divergent organisms.

An elegant and direct way of exploring the mechanisms of phenotypic evolution and potential processes of species diversification is to compare the expression of candidate genes across a clade of phylogenetically related but morphologically divergent species. Abzahnov and colleagues (2004) examined the expression patterns of several growth factors in six species of Darwin's finches to test their potential role in beak morphogenesis and evolution. These six bird species, which display considerable variation in beak size and shape, are the sharp-beaked finch (*Geospiza difficilis*), the small ground finch (*Geospiza Fuliginosa*), the medium ground finch (*Geospiza Fortis*), the large ground finch (*Geospiza magnirostoris*), the cactus finch (*Geospiza scandens*), and the large cactus finch (*Geospiza conirostris*).

Fibroblast growth factor 8 (*FGF8*) and Sonic hedgehog (*Shh*), which are believed to pattern the mesenchyme and drive cartilage growth (Abzhanov and Tabin 2004), did not differ in expression between the finch embryos. However, the expression of three bone morphogenetic protein (*Bmp2, Bmp4,* and *Bmp7*) genes was found to correlate with beak morphometrics across the species. Specifically, transcript levels of *Bmp4* were found to correlate with the shape of developing beaks by increasing beaks' depth and breadth (Abzhanov et al. 2004). Phenotypic evolution, such as the emergence of different beak shapes, may hence be facilitated by changes in the expression levels of specific developmental genes, rather than by the origin of entirely novel genes or pathways.

Active manipulation of candidate gene action in nonmodel avian species is still in its infancy, but the recent developmental studies of beak morphology again illustrate the ways in which experiments on model and nonmodel species can illuminate each other. Indeed, to understand the species-specific mechanisms responsible for generating differences in beak morphology, these studies have again relied on the chicken for providing a critical interspecific comparison, as well as a technically easier means of gene manipulation.

Wu and coworkers (2004) identified the developmental gene *Bmp4* as differentially expressed in the beaks of chicken and ducks, and hence potentially driving differences in beak shape. They then confirmed that *Bmp4* activity was associated with larger beak sizes by injecting *Bmp4* into an avian retroviral vector (box 1), subsequently used to infect the beak prominences of chicken embryos. Conversely, when they infected developing chick embryos with a retrovirus vector containing the antagonist of *Bmp4,* they observed the expected reduction of beak size.

In a similar way, Abzhanov and colleagues (2006) tested whether calmodulin (CaM) pathway activation drives beak length development in Darwin's finches by inserting a constitutively activated form of the CaM kinase kinase (*CaMKII*) into an avian retroviral vector, and subsequently injecting it into the frontonasal processes of chicken embryos. The expression of the activated form of *CaMKII* resulted in a significant elongation of the chicken beak, without affecting the other beak axes (i.e., width and depth), confirming the independent regulation of axes of beak shape and the potential for using chicken in tests of gene function to complement evolutionary studies of wild organisms (Abzhanov et al. 2006).

Genomewide transcriptional profiling

Quantifying differences in the amount of transcripts between two phenotypes or treatment groups is interesting not only for testing for potential effects of candidate genes but also for identifying novel genes or pathways associated with phenotypic differences. Furthermore, because variation in levels of transcription can mediate adaptive evolution (Enard et al. 2002), quantitative changes in mRNA (messenger ribonucleic acid) can also help us understand how populations may evolve in response to selection. Clayton and Huecas (1990) were the first to attempt to estimate the number of genes expressed in a wild bird, the canary. They specifically isolated 5000 or so rare RNAs transcribed in the forebrain, the region of song production, but not in the rest of the brain.

Complementary DNA libraries and expressed sequence tags. A critical step toward sequencing of the chicken genome was high-throughput DNA sequencing of expressed sequence tags (ESTs) from dozens of tissue-specific cDNA libraries generated from several international projects (box 1; for a review, see Burt [2005]). This feat has advanced the chicken to 14th place (with 599,330 ESTs) among all model organisms represented in the dbEST division of GenBank. Recently, a multi-institutional initiative was launched to create a complete EST database for genes expressed in the zebra finch brain (*http://titan.biotec.uiuc.edu/songbird/*), and an initial version of a database containing the cDNA clone information was made publicly accessible as of 2002. This initiative also facilitates the use of microarrays to link transcript profiles with functional and developmental states of the zebra finch. Because of its importance to evolutionary and ecological studies, the zebra finch genome project opens up many new and exciting avenues of evolutionary research, both at the intra- and interspecific levels.

For example, the first large-scale analyses of protein evolution in birds have capitalized on comparisons between chicken and the zebra finch, particularly at genes produced by the Neurogenomics Initiative (Mank et al. 2007). These studies suggest a range of degrees of adaptive evolution in birds and show that the "fast-Z effect"—the hypothesis that recessive but beneficial mutations will accumulate faster on the hemizygous Z chromosome than on autosomes—is alive and well in avian genome evolution.

Several cDNA libraries have already been made for other bird species, such as for turkeys (Smith et al. 2000, Dranchak et al. 2003, Chaves et al. 2005), zebra finches (Wade et al. 2005), Japanese quails (*Coturnix coturnix japonica*) (Mott and Ivarie

2004), house finches (Wang et al. 2006b), and domestic mallards (*Anas platyrhynchos*) (Xia et al. 2007). Partial sequencing of cDNA libraries allows the development of ESTs, which can then be used in comparative genome analyses of birds. Sequencing ESTs in one species can also provide means of exploring sequence variation in other bird species. For example, Smith and colleagues (2000) developed primers for 21 ESTs using the cDNA of turkey pituitary, which were then used in chicken, guinea fowl (*Numida meleagris*), pigeon (*Columba domestica*), and quail.

The specific location of the expression of transcripts can be visualized by *in situ* hybridizations (box 1). This technique has been used extensively in bird species of commercial interest, particularly in developmental contexts (e.g., Antin et al. 2007, Voigt et al. 2007). For example, the GEISHA project (*Gallus* Expression *in situ* Hybridization Analysis; *http://geisha.arizona.edu/geisha/*) centralizes all the expression data obtained from *in situ* hybridizations of the chicken embryo into a single database that can be queried for different genes and developmental stages (Antin et al. 2007), thereby supplying useful data for parallel studies in other birds (Burt 2005).

MicroRNAs. Animal, plant, and viral genomes encode small, noncoding RNAs that regulate gene function by affecting stability or translational efficiency of target mRNAs. One class of small RNAs, the microRNAs (miRNAs), is becoming widely appreciated as a pivotal regulator of gene expression. In animal cells, miRNAs suppress gene function primarily by blocking translation of the mRNA. Each miRNA could potentially target a large number of mRNAs, and it is thought that a large portion of the transcriptome is regulated by miRNAs (Lim et al. 2005).

There is strong phylogenetic conservation of the sequence of these molecules, although several studies have predicted or identified species-specific miRNAs (Ambros 2004, Bentwich et al. 2005). Chicken miRNAs, 122 homologs, were identified in the analysis of the chicken genome (ICGSC 2004) by comparison with other species. In a detailed study of the chicken transcriptome, 23 evolutionarily conserved miRNAs were found in chicken EST databases (Hubbard et al. 2005), and *in situ* hybridization studies with chick embryos confirmed the expression and, in some instances, the tissue-specific expression of several miRNAs (Darnell et al. 2006). A more in-depth analysis of miRNA expression in chick development has been presented (Darnell et al. 2006), and a comparative analysis with mammalian developmental profiles documents some similarities as well as differences in expression profiles. Studies on microRNAs in comparative bird development will help to refine our understanding of these molecules.

Subtractive suppression hybridization and macroarrays. Wang and colleagues (2006b) used the subtractive suppression hybridization (SSH) technique (box 1) to identify house finch genes that responded to infection by a bacterium, *Mycoplasma gallisepticum* (MG), and that might therefore undergo MG-driven natural selection. They constructed two subtractive cDNA libraries, one made of clones found to be up-regulated in the spleen of MG-infected birds relative to controls, and a second of clones found to be up-regulated in the spleen of healthy birds relative to MG-infected ones. A total of 220 cDNA clones, consisting of 34 genes with known homologues and novel transcripts, were found to be qualitatively either up-regulated or down-regulated by high-density filter hybridization. Putative gene expression changes were subsequently confirmed by high-throughout reverse northern blot hybridizations (see "Macroarrays" in box 1). Sequencing and BLAST analysis of target clones identified heat shock protein 90, MHC II-associated invariant chain (CD74), TIM1 (T-cell immunoglobulin mucin 1), and granzyme A, among several others, as genes that were strongly differentially expressed between MG-infected and healthy house finches (Wang et al. 2006b). So far, this is the only study employing this particular approach to transcriptional profiling in birds.

Microarrays. The initial chicken lymphoid cDNA microarrays (box 1) provided the first glimpse of global gene expression in the chicken's immune system (Morgan et al. 2001, Neiman et al. 2001, Cui et al. 2004), and tissue-specific DNA microarrays have since been developed for transcriptional profiling in liver, pineal gland, retina, metabolic and somatic, neuroendocrine and reproductive systems, macrophages, and intestinal lymphocytes (for a review, see Cogburn et al. [2007]). Genomewide microarray investigations have been recently employed to examine differences in the transcription levels of sex-linked genes and sex-chromosome dosage compensation in chicken (Scholz et al. 2006, Ellegren et al. 2007, Melamed and Arnold 2007). That the expression levels of Z-linked genes are higher overall in male chicken embryos than in female ones, both in somatic tissues and gonads, suggests that chromosome-wide adjustment of sex-linked gene expression may be less effective in birds than in mammals, or even nonexistent (Ellegren et al. 2007, Itoh et al. 2007).

Complementary DNA microarrays can allow the identification of previously unknown genes and pathways whose expression correlate with a specific phenotype. Abzhanov and colleagues (2006) designed a microarray using a cDNA library constructed from RNA isolated from the embryonic frontonasal processes of one of Darwin's finches, *G. fortis.* The goal of the study was to identify new genetic pathways involved in Darwin's finch beak morphogenesis by comparing gene expression profiles between species belonging to a monophyletic group but exhibiting distinct beak morphology—that is, variation in beak depth, width, or length (or some combination of these). Individual microarray slides were hybridized with embryonic frontonasal primordia RNA from five species of Darwin's finches and a common reference of several individuals belonging to a basal species, *G. difficilis.* These microarray analyses identified the calmodulin pathway in the promotion of beak elongation, adding to the BMP4 pathway as a driver of finch beak evolution. The implication

of the calmodulin pathway in beak formation was subsequently validated *in situ* by comparing Darwin finch embryos hybridized with a chick antisense riboprobe (box 1) of the gene *CaM* (figure 2). The multiple regulatory pathways responsible for beak growth along different morphological axes implied by these studies were hypothesized to allow natural selection to generate diversity in one axis independently of the other, thereby magnifying the range of potential variation in beak morphology.

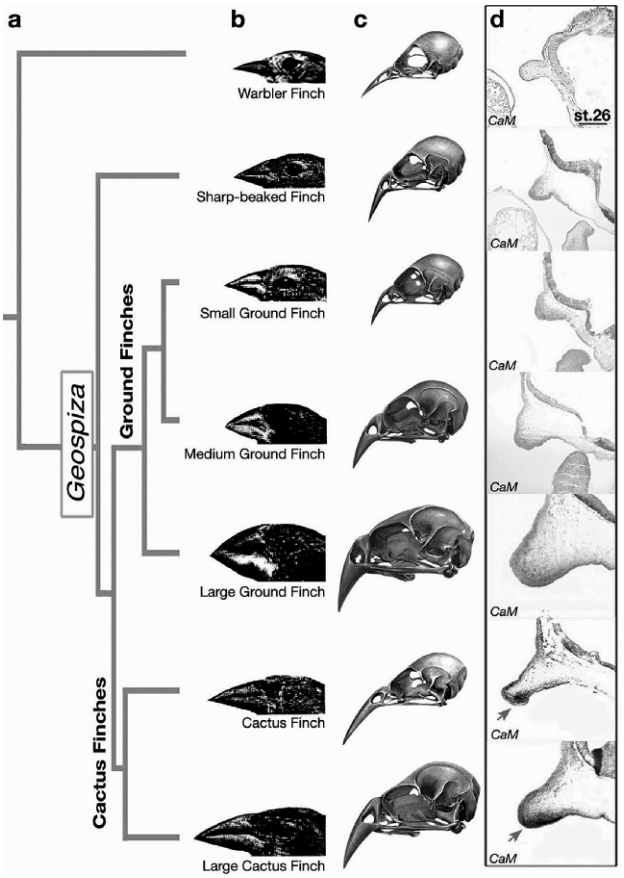

Figure 2. (a–c) Differences in beak morphologies among the Darwin's finches. The sharp-beaked ground finch (*Geospiza difficilis*) is the basal species of the genus* Geospiza*; the cactus finch (*Geospiza scandens*) and the large cactus finch (*Geospiza conirostris*) comprise the cactus finches; and the small ground finch (*Geospiza fuliginosa*), the medium ground finch (*Geospiza fortis*), and the large ground finch (*Geospiza magnirostris*) comprise the ground finches. (d) Darwin's finch embryos were harvested at stage 26 and hybridized* in situ *with a chick probe of calmodulin (*CaM*). CaM was expressed at its highest levels in the distal-ventral mesenchyme of the upper beak prominence of the large cactus finch, and at decreasingly lower levels in the cactus finch and in the large ground finch and the medium ground finch. Source: Abzhanov and colleagues (2006), with permission.

Conclusions

The avian genome is ripe for genomic analysis, and the frenzy of large-scale analyses, at both the molecular and the computational levels, has begun. The streamlined genomes of birds will provide a simplified template with which to understand the determinants and evolution of vertebrate gene regulation. We wait with excitement to see how the density of global expression patterns and transcription factor binding sites in birds differ from those of other vertebrates whose genomes harbor vastly greater amounts of nongenic DNA. Micro- and macroarray studies have already yielded insight into avian ecology, in particular in the area of the immunogenetics of pathogen infection. Novel deep-sequencing strategies promise to complete the inventory of functional elements of avian genomes, in particular the microRNAs and other noncoding RNAs. Full genome sequences and large-scale sequence scanning in nonmodel birds and reptiles will open the way to truly comparative genomics, unleash vast numbers of loci for phylogenetic analysis, and permit the first global analyses of protein and gene evolution. All of these emerging technologies stand to benefit from application to diverse avian and other nonmodel species in novel evolutionary and ecological genomic investigations.

Acknowledgments

We are grateful to Arkhat Abzhanov and three anonymous referees for their valuable comments on a previous version of the manuscript. We also thank Hans Ellegren for sending us his presubmission manuscript. We apologize to those authors whose work could not be cited because of space constraints.

References cited

Abzhanov A, Tabin CJ. 2004. *Shh* and *Fgf8* act synergistically to drive cartilage outgrowth during cranial development. Developmental Biology 273: 134–148.

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ. 2004. *Bmp4* and morphological variation of beaks in Darwin's finches. Science 305: 1462–1465.

Abzhanov A, Kuo WP, Hartmann C, Grant BR, Grant PR, Tabin CJ. 2006. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. Nature 442: 563–567.

Ambros V. 2004. The function of animal microRNAs. Nature 431: 350–355.

Antin PB, Kaur S, Stanislaw S, Davey S, Konieczka JH, Yatskievych TA, Darnell DK. 2007. *Gallus* expression *in situ* hybridization analysis: A chicken embryo gene expression database. Poultry Science 86: 1472–1477.

Axelsson E, Webster MT, Smith NG, Burt DW, Ellegren H. 2005. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. Genome Research 15: 120–125.

Backström N, Fagerberg S, Ellegren H. 2008. A gene-based set of reference markers evenly spread across the avian genome. Molecular Ecology 17: 964–980.

Bensch S, Irwin DE, Irwin JH, Kvist L, Akesson S. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. Molecular Ecology 15: 161–171.

Bentley DR. 2006. Whole-genome re-sequencing. Current Opinion in Genetics and Development 16: 545–552.

Bentwich I, et al. 2005. Identification of hundreds of conserved and nonconserved human microRNAs. Nature Genetics 37: 766–770.

Burt DW. 2002. Origin and evolution of avian microchromosomes. Cytogenetics and Genome Research 96: 97–112.

———. 2005. Chicken genome: Current status and future opportunities. Genome Research 15: 1692–1698.
Chaves LD, Rowe JA, Reed KM. 2005. Survey of a cDNA library from the turkey (*Meleagris gallopavo*). Genome 48: 12–17.
Clayton DF. 2004. Songbird genomics: Methods, mechanisms, opportunities, and pitfalls. Pages 45–60 in Zeigler HP, Marler P, eds. Behavioral Neurobiology of Birdsong, vol. 1016. New York: New York Academy of Sciences.
Clayton DF, Huecas M. 1990. Forebrain-enriched RNAs of the canary: A population analysis using hybridization kinetics. Molecular Brain Research 7: 23–30.
Cogburn LA, Porter TE, Duclos MJ, Simon J, Burgess SC, Zhu JJ, Cheng HH, Dodgson JB, Burnside J. 2007. Functional genomics of the chicken—a model organism. Poultry Science 86: 2059–2094.
Cui J, Sofer L, Cloud SS, Burnside J. 2004. Patterns of gene expression in the developing chick thymus. Developmental Dynamics 229: 480–488.
Dall SRX, Houston AI, McNamara JM. 2004. The behavioural ecology of personality: Consistent individual differences from an adaptive perspective. Ecology Letters 7: 734–739.
Darnell DK, Kaur S, Stanislaw S, Konieczka JK, Yatskievych TA, Antin PB. 2006. MicroRNA expression during chick embryo development. Developmental Dynamics 235: 3156–3165.
Dawson DA, Akesson M, Burke T, Pemberton JM, Slate J, Hansson B. 2007. Gene order and recombination rate in homologous chromosome regions of the chicken and a passerine bird. Molecular Biology and Evolution 24: 1537–1552.
Dranchak PK, Chaves LD, Rowe JA, Reed KM. 2003. Turkey microsatellite loci from an embryonic cDNA library. Poultry Science 82: 526–531.
Edwards SV, Jennings WB, Shedlock AM. 2005. Phylogenetics of modern birds in the era of genomics. Proceedings of the Royal Society B 272: 979–992.
Ellegren H. 2005. The avian genome uncovered. Trends in Ecology and Evolution 20: 180–186.
———. 2007. Molecular evolutionary genomics of birds. Cytogenetic and Genome Research 117: 120–130.
Ellegren H, Hultin-Rosenberg L, Brunstrom B, Dencker L, Kultima K, Scholz B. 2007. Faced with inequality: Chicken do not have a general dosage compensation of sex-linked genes. BMC Biology 5: 40.
Enard W, et al. 2002. Intra- and interspecific variation in primate gene expression patterns. Science 296: 340–343.
Fidler AE, van Oers K, Drent PJ, Kuhn S, Mueller JC, Kempenaers B. 2007. *Drd4* gene polymorphisms are associated with personality variation in a passerine bird. Proceedings of the Royal Society B 274: 1685–1691. doi:10.1098/rspb.2007.0337
Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbel JO, Emanuelsson O, Zhang ZD, Weissman S, Snyder M. 2007. What is a gene, post-ENCODE? History and updated definition. Genome Research 17: 669–681.
Gregory TR. 2007. Animal Genome Size Database. (6 June 2008; *www.genomesize.com*)
Hardison RC. 2003. Comparative genomics. PloS Biology 1: 156–160.
Hedges SB. 2002. The origin and evolution of model organisms. Nature Reviews Genetics 3: 838–849.
Helfer G, Fidler AE, Vallone D, Foulkes NS, Brandstaetter R. 2006. Molecular analysis of clock gene expression in the avian brain. Chronobiology International 23: 113–127.
Henikoff S. 2007. ENCODE and our very busy genome. Nature Genetics 39: 817–818.
Hubbard SJ, et al. 2005. Transcriptome analysis for the chicken based on 19,626 finished cDNA sequences and 485,337 expressed sequence tags. Genome Research 15: 174–183.
Hughes AL. 2000. Adaptive Evolution of Genes and Genomes. Oxford (United Kingdom): Oxford University Press.
Hughes AL, Piontkivska H. 2005. DNA repeat arrays in chicken and human genomes and the adaptive evolution of avian genome size. BMC Evolutionary Biology 5: 12.
[ICGSC] International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432: 695–716.
Iglesias GM, Soria LA, Goto RM, Jar AM, Miquel MC, Lopez OJ, Miller MM. 2003. Genotypic variability at the major histocompatibility complex (B and Rfp-Y) in Camperos broiler chickens. Animal Genetics 34: 88–95.
Itoh Y, et al. 2007. Dosage compensation is less effective in birds than in mammals. Journal of Biology 6: 2.
Jennings WB, Edwards SV. 2005. Speciational history of Australian grass finches (Poephila) inferred from thirty gene trees. Evolution 59: 2033–2047.
Kaufman J, Milne S, Gobel TWF, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S. 1999. The chicken B locus is a minimal essential major histocompatibility complex. Nature 401: 923–925.
Keeling L, Anderson L, Schutz KE, Kerje S, Fredriksson R, Carlborg O, Cornwallis CK, Pizzari T, Jensen P. 2004. Chicken genomics: Feather-pecking and victim pigmentation. Nature 431: 645–646.
Kluger AN, Siegfried Z, Ebstein RP. 2002. A meta-analysis of the association between *DRD4* polymorphism and novelty seeking. Molecular Psychiatry 7: 712–717.
LePage KT, Miller MM, Briles WE, Taylor RL Jr. 2000. Rfp-Y genotype affects the fate of Rous sarcomas in B2B5 chickens. Immunogenetics 51: 751–754.
Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433: 769–773.
Luo MZ, et al. 2006. Utilization of a zebra finch BAC library to determine the structure of an avian androgen receptor genomic region. Genomics 87: 181–190.
Mank JE, Axelsson E, Ellegren H. 2007. Fast-X on the Z: Rapid evolution of sex-linked genes in birds. Genome Research 17: 618–624.
Margulies M, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376–380.
Melamed E, Arnold AP. 2007. Regional differences in dosage compensation on the chicken Z chromosome. Genome Biology 8: R202.
Miyake T, Amemiya CT. 2004. BAC libraries and comparative genomics of aquatic chordate species. Comparative Biochemistry and Physiology C: Toxicology and Pharmacology 138: 233–244.
Morgan RW, Sofer L, Anderson AS, Bernberg EL, Cui J, Burnside J. 2001. Induction of host gene expression following infection of chicken embryo fibroblasts with oncogenic Marek's disease virus. Journal of Virology 75: 533–539.
Mott IW, Ivarie R. 2004. cDNA array analysis of Japanese quail lines divergently selected for four-week body weight. Poultry Science 83: 1524–1529.
Mundy NI. 2005. A window on the genetics of evolution: *MC1R* and plumage colouration in birds. Proceedings of the Royal Society B 272: 1633–1640.
Nadeau NJ, Burke T, Mundy NI. 2007. Evolution of an avian pigmentation gene correlates with a measure of sexual selection. Proceedings of the Royal Society B 274: 1807–1813.
Neiman PE, Ruddell A, Jasoni C, Loring G, Thomas SJ, Brandvold KA, Lee RM, Burnside J, Delrow J. 2001. Analysis of gene expression during *myc* oncogene-induced lymphomagenesis in the bursa of Fabricius. Proceedings of the National Academy of Sciences 98: 6378–6383.
Organ CL, Shedlock AM, Meade A, Pagel M, Edwards SV. 2007. Origin of avian genome size and structure in non-avian dinosaurs. Nature 446: 180–184.
Primmer CR, Raudsepp T, Chowdhary BP, Moller AR, Ellegren H. 1997. Low frequency of microsatellites in the avian genome. Genome Research 7: 471–482.
Richardson DS, Komdeur J, Burke T, von Schantz T. 2005. MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. Proceedings of the Royal Society B 272: 759–767.
Romanov MN, Dodgson JB. 2006. Cross-species overgo hybridization and comparative physical mapping within avian genomes. Animal Genetics 37: 397–399.
Romanov MN, Koriabine M, Nefedov M, de Jong PJ, Ryder OA. 2006. Construction of a California condor BAC library and first-generation chicken-condor comparative physical map as an endangered species conservation genomics resource. Genomics 88: 711–718.

Scholz B, Kultima K, Mattsson A, Axelsson J, Brunstrom B, Halldin K, Stigson M, Dencker L. 2006. Sex-dependent gene expression in early brain development of chicken embryos. BMC Neuroscience 7: 12.

Shaw I, Powell TJ, Marston DA, Baker K, van Hateren A, Riegert P, Wiles MV, Milne S, Beck S, Kaufman J. 2007. Different evolutionary histories of the two classical class I genes *BF1* and *BF2* illustrate drift and selection within the stable MHC haplotypes of chickens. Journal of Immunology 178: 5744–5752.

Shedlock AM, Botka CW, Zhao S, Shetty J, Zhang T, Liu JS, Deschavanne PJ, Edwards SV. 2007. Phylogenomics of nonavian reptiles and the structure of the ancestral amniote genome. Proceedings of the National Academy of Sciences 104: 2767–2772.

Shiina T, Briles WE, Goto RM, Hosomichi K, Yanagiya K, Shimizu S, Inoko H, Miller MM. 2007. Extended gene map reveals tripartite motif, C-type lectin, and Ig superfamily type genes within a subregion of the chicken *MHC-B* affecting infectious disease. Journal of Immunology 178: 7162–7172.

Sih A, Bell A, Johnson JC. 2004. Behavioral syndromes: An ecological and evolutionary overview. Trends in Ecology and Evolution 19: 372–378.

Smith E, Shi L, Drummond P, Rodriguez L, Hamilton R, Powell E, Nahashon S, Ramlal S, Smith G, Foster J. 2000. Development and characterization of expressed sequence tags for the turkey (*Meleagris gallopavo*) genome and comparative sequence analysis with other birds. Animal Genetics 31: 62–67.

Strand T, Westerdahl H, Höglund J, Alatalo RV, Siitari H. 2007. The *Mhc* class II of the black grouse (*Tetrao tetrix*) consists of low numbers of *B* and *Y* genes with variable diversity and expression. Immunogenetics 59: 725–734.

van Oers K, de Jong G, van Noordwijk AJ, Kempenaers B, Drent PJ. 2005. Contribution of genetics to the study of animal personalities: A review of case studies. Behaviour 142: 1185–1206.

Voigt C, Ball GF, Balthazart J. 2007. Neuroanatomical specificity of sex differences in expression of aromatase mRNA in the quail brain. Journal of Chemical Neuroanatomy 33: 75–86.

Wade J, Peabody C, Coussens P, Tempelman RJ, Clayton DF, Liu L, Arnold AP, Agate R. 2005. A cDNA microarray from the telencephalon of juvenile male and female zebra finches. Journal of Neuroscience Methods 138: 199–206.

Waltari E, Edwards SV. 2002. Evolutionary dynamics of intron size, genome size, and physiological correlates in archosaurs. American Naturalist 160: 539–552.

Wang ZS, Miyake T, Edwards SV, Amemiya CT. 2006a. Tuatara (*Sphenodon*) genomics: BAC library construction, sequence survey, and application to the *DMRT* gene family. Journal of Heredity 97: 541–548.

Wang ZS, Farmer K, Hill GE, Edwards SV. 2006b. A cDNA macroarray approach to parasite-induced gene expression changes in a songbird host: Genetic response of house finches to experimental infection by *Mycoplasma gallisepticum*. Molecular Ecology 15: 1263–1273.

Webster MT, Axelsson E, Ellegren H. 2006. Strong regional biases in nucleotide substitution in the chicken genome. Molecular Biology and Evolution 23: 1203–1216.

Wu P, Jiang TX, Suksaweang S, Widelitz RB, Chuong CM. 2004. Molecular shaping of the beak. Science 305: 1465–1466.

Wyles JS, Kunkel JG, Wilson AC. 1983. Birds, behavior, and anatomical evolution. Proceedings of the National Academy of Sciences 80: 4394–4397.

Xia JG, Radford C, Guo XX, Magor KE. 2007. Immune gene discovery by expressed sequence tag analysis of spleen in the duck (*Anas platyrhynchos*). Developmental and Comparative Immunology 31: 272–285.

doi:10.1641/B580706
Include this information when citing this material.

Are We Consuming Too Much?
The answer seems obvious. But it's not. Paul Ehrlich, Kenneth Arrow, and nine other brilliant minds argue that we're worrying too much about how much we consume and too little about how to invest.

Why Aren't Fish Populations Recovering?
Evidence is mounting that fish populations won't necessarily recover even if fishing stops. Fishing may be such a powerful evolutionary force that we are running up a Darwinian debt for future generations.

What Makes Environmental Treaties Work?
If the Kyoto Protocol fails, will it be because George Bush cares more about the oil industry than global warming? Or because it was a badly thought-out treaty?

Are Linguistic and Biological Diversity Linked?
Ten out of the top 12 countries containing the world's highest biodiversity overlap with the top 25 countries for number of endemic languages.